Toward Guaranteed Illumination Models for Non-Convex Objects

Yuqian Zhang, Cun Mu, Han-Wen Kuo, John Wright
Columbia University in the City of New York
{yz2409, cm3052, hk2673}@columbia.edu, johnwright@ee.columbia.edu

Abstract

Illumination variation remains a central challenge in object detection and recognition. Existing analyses of illumination variation typically pertain to convex, Lambertian objects, and guarantee quality of approximation in an average case sense. We show that it is possible to build models for the set of images across illumination variation with worst-case performance guarantees, for nonconvex Lambertian objects. Namely, a natural verification test based on the distance to the model guarantees to accept any image which can be sufficiently well-approximated by an image of the object under some admissible lighting condition, and guarantees to reject any image that does not have a sufficiently good approximation. These models are generated by sampling illumination directions with sufficient density, which follows from a new perturbation bound for directional illuminated images in the Lambertian model. As the number of such images required for guaranteed verification may be large, we introduce a new formulation for cone preserving dimensionality reduction, which leverages tools from sparse and low-rank decomposition to reduce the complexity, while controlling the approximation error with respect to the original model.[1]

1. Introduction

Illumination variation remains a central challenge in object detection and recognition. Changes in lighting can dramatically change the appearance of the object, rendering simple pattern recognition techniques such as nearest neighbor ineffective. Various approaches have been proposed to mitigate this problem, for example using nonlinear features based on gradient orientation [13], using quotient images [18] or total variation regularization [6]. These approaches are often effective in practice, but can break down under extreme illumination. Moreover, because of the nonlinearity of the feature extraction step, clearly characterizing their domain of applicability is challenging.

An alternative approach is to attempt to explicitly characterize the set of images of the object generated under

varying lighting. The seminal work [2] argues that images of a given object with fixed pose and varying illumination should lie near a convex cone in the high-dimensional image space. Many subsequent works have attempted to capture the gross structure of this cone using low-dimensional convex cone or linear subspace models [1, 15]. These models have been used for recognition in many subsequent works [9, 20, 21, 19], and have been extended in a number of directions [8, 16]. The promise of subspace or cone models, compared to feature-based approaches described above, is that, by reasoning carefully about the image formation process, it might be possible to guarantee to well-approximate all images of the object under clearly delineated conditions.

It is worth asking then, what approximation guarantees do current results afford us? For convex, Lambertian objects, by an elegant interpretation of the Lambertian reflectance as spherical convolution [1, 15], people showed that for uniformly random directional lightings, a nine dimensional spherical harmonic approximation captures on average about 98% of the energy [1, 8]. In this work, we ask whether it is possible to build models that (i) guarantee robustness to *worst* case lighting, over some clearly specified class of admissible lighting conditions, and (ii) work even for *nonconvex* objects, whose cast shadows create moving boundaries and render low-dimensional linear model ineffective, and (iii) have low storage and computational complexity? We study these questions in the context of a model problem in object instance verification, in which one is given an object $\mathcal{O}$ at a fixed pose, and ask whether the input image is an image of this object under some valid illumination condition. We develop rigorous guarantees for this problem, for general (including nonconvex) Lambertian objects. Our results show how to build a model that guarantees to accept every image that can be interpreted as an image of the object under some lighting condition, and to reject every image that is sufficiently dissimilar to all images of the object under valid lighting conditions.

Similar to [12, 14, 19], our work approximates the illumination cone with a conic combination of certain images on its boundary. Empirically, such constructions may require relatively large numbers of images [14], but no rigorous results of this nature are currently known. To address

---

this problem, we start from the goal of producing a sufficiently accurate representation of the illumination cone, and derive, in terms of the properties of the object and the scene, sufficient sampling densities for this goal to be met. Our bounds depend on the level of *ambient illumination*, and a notion of *convexity defect* of the object. They make precise the intuitions that: (i) it is more difficult to operate in low-light scenarios and (ii) nonconvex objects are more challenging than convex objects.

The number of images required to guarantee performance can be large. To address this problem, we introduce a new approach to cone preserving complexity reduction. This approach leverages tools from convex programming – in particular, sparse and low-rank decomposition [4, 5] – but introduces a new constrained formulation which guarantees that the conic hull of the output will well-approximate the conic hull of the input. The low-rank and sparse decomposition leverages our qualitative understanding of the physical properties of images (low-dimensionality, sparsity of cast shadows) [1, 15, 21, 4, 22], while the constraint ensures that the output of this algorithm is always a good approximation to the target cone. Empirically, we find that the output is often of much lower complexity than the input. This suggests a methodology for building instance verifiers that are both robust to worst case illumination, and computationally efficient. Numerical experiments illustrate our bounds and their potential for worst case verification. All mathematical claims are proved in the appendix.

2. Problem Formulation and Methodology

Cone Models for Illumination. We consider images of size $w \times h$, and let $m = wh$, each image can be treated as a vector $\boldsymbol{y} \in \mathbb{R}^m$. We are interested in the set of images of an object $\mathcal{O}$ that can be generated under distant illumination. These images form a subset $C_0 \subseteq \mathbb{R}^m$. Each distant illumination can be identified with a nonnegative function $f : \mathbb{S}^2 \to \mathbb{R}_+$, whose value $f(\boldsymbol{u})$ is the intensity of light from direction $\boldsymbol{u}$. We use the notation $\mathcal{F}$ for the set of nonnegative, Riemann integrable functions on $\mathbb{S}^2$. Mathematically, $\mathcal{F}$ is a convex cone: sums of nonnegative, integrable functions are again nonnegative and integrable.

We assume a linear sensor response: the image is a linear function of the incident irradiance. By linearity of light transport and linearity of the sensor response, the observed image $\boldsymbol{y} \in \mathbb{R}^m$ is a linear function $\boldsymbol{y}[f]$ of the illumination f: if the object is subjected to the superposition $f = f_1 + f_2$ of two illuminations f_1 and f_2, we observe $\boldsymbol{y}[f_1 + f_2] = \boldsymbol{y}[f_1] + \boldsymbol{y}[f_2]$. Since f resides in the convex cone, the set $C_0 \doteq \boldsymbol{y}[\mathcal{F}] \subset \mathbb{R}^m$ of possible images is also a convex cone. Note, however, that the fact that C_0 is a convex cone holds under very mild assumptions.

The detailed properties of C_0 were first studied in [2], and a great deal of subsequent work has been devoted to
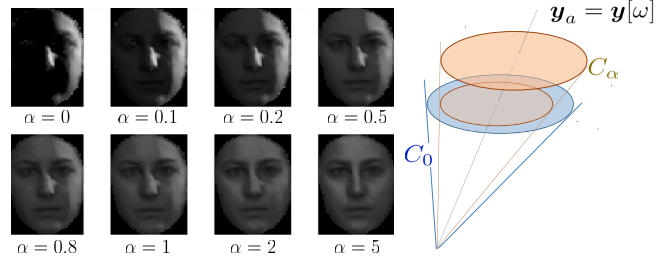

Figure 1. **Ambient level α.** Left: typical images from the cone C_α, for ambient levels $\alpha = 0$ up to $\alpha = 5$. In each example f_d is an extreme directional illumination. Images rendered from [17]. Right: illumination cones C_α with varying ambient level α.

understanding its properties [15, 1, 8]. Most of this body of work has been devoted to simple, analytically tractable models such as convex, Lambertian objects. For such objects, interesting statements can be made about the gross shape of C_0.

The cone C_0 can be interpreted as the set of all images of the object under different distant lighting conditions. Intuitively speaking, we expect the problem of representing images $\boldsymbol{y}$ under different illuminations to be more challenging when the light has a stronger directional component. To capture the relative contribution of directional and ambient components of light, we introduce a family of function classes $\mathcal{F}_\alpha$, indexed by parameter $\alpha \in [0, \infty)$. Illuminations in $\mathcal{F}_\alpha$ consist of an ambient component $\alpha\omega$, where $\omega(\boldsymbol{u}) = 1/\text{area}(\mathbb{S}^2)$ is the constant function on the sphere, and an arbitrary (possibly directional) component f_d:

$$\mathcal{F}_\alpha = \left\{ f_d + \alpha\omega \mid f_d \in \mathcal{F},\ \|f_d\|_{L_1} \le 1 \right\}, \qquad (1)$$

For each ambient level α, we have a cone

$$C_\alpha \doteq \left\{ t\boldsymbol{y}[f] \mid t \ge 0, f \in \mathcal{F}_\alpha \right\} \qquad (2)$$

For any $\alpha \le \alpha'$, $C_{\alpha'} \subseteq C_\alpha$. In this sense, the choice of α induces a tradeoff: as α becomes smaller, C_α becomes more complicated to compute with, but can represent broader illumination conditions. Our complexity bounds in Section 3 will make this intuition precise. Figure 1 shows rendered images of a face under various ambient levels $\alpha \ge 0$. Our methodology is compatible with any choice of $\alpha > 0$.

Verification using Convex Cones. Our methodology asks the system designer to select a target level of ambient illumination α, and hence choose a target cone $C = C_\alpha$. At test time, we are given a new input image $\boldsymbol{y} \in \mathbb{R}^m$. The verification problem asks us to decide if $\boldsymbol{y}$ could be an image of object $\mathcal{O}$: *Is $\boldsymbol{y}$ an element of C?* Or, if $\boldsymbol{y}$ is subject to noise, *Is $\boldsymbol{y}$ sufficiently close to C?* The distance from $\boldsymbol{y}$ to C in ℓ^2-norm is $d(\boldsymbol{y}, C) \doteq \inf \left\{ \|\boldsymbol{y} - \boldsymbol{y}'\|_2 \mid \boldsymbol{y}' \in C \right\}$. Any cone C is nonnegatively homogeneous: if $\boldsymbol{z} \in C$, $t\boldsymbol{z} \in C$ for all $t \ge 0$. To obtain a scale invariant criterion, we can work with the angle: $\angle(\boldsymbol{y}, C) \doteq \text{asin}(d(\boldsymbol{y}, C)/\|\boldsymbol{y}\|_2)$, giving the following simple, natural test for object detection:
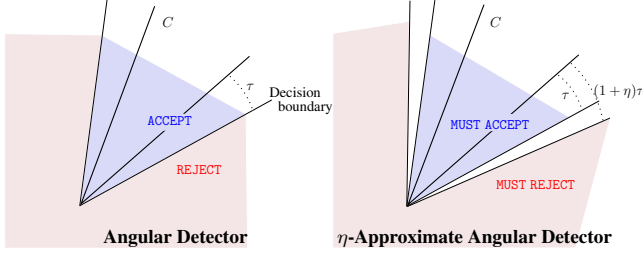
Figure 2. Two detection rules. The angular detector accepts points based on their angle with the cone C. An approximate angular detector guarantees to accept any point within angle τ of C, and to reject any point with angle greater than $(1+\eta)\tau$. In the intermediate region (white) there are no restrictions on its behavior.

Definition 2.1. *The angular detector (AD)* $\mathfrak{D}_\tau^C : \mathbb{R}_+^m \to \{\mathrm{ACCEPT}, \mathrm{REJECT}\}$ *with threshold* τ *is the decision rule*

$$\mathfrak{D}_\tau^C(\boldsymbol{y}) = \begin{cases} \mathrm{ACCEPT} & \angle(\boldsymbol{y},\, C) \leq \tau, \\ \mathrm{REJECT} & \angle(\boldsymbol{y},\, C) > \tau. \end{cases} \quad (3)$$

If C is a polyhedral cone, the decision rule (3) can be implemented using nonnegative least squares. This is efficient if the number n of extreme rays of C is small. If $\mathcal{O}$ is a convex polyhedron with only a few vertices, this is the case. However, in general, the number of extreme rays in a $\mathcal{V}$(vertex)-description can be large or even unbounded.[2] One remedy is to relax the definition slightly:

Definition 2.2. *The* η*-approximate angular detector* $(\eta-AAD)$ $\widehat{\mathbb{D}}_{\tau,\eta}^C : \mathbb{R}_+^m \to \{\mathrm{ACCEPT}, \mathrm{REJECT}\}$ *satisfies*

$$\widehat{\mathbb{D}}_{\tau,\eta}^C(\boldsymbol{y}) = \begin{cases} \mathrm{ACCEPT} & \angle(\boldsymbol{y},\, C) \leq \tau, \\ \mathrm{REJECT} & \angle(\boldsymbol{y},\, C) > (1+\eta)\,\tau. \end{cases} \quad (4)$$

Figure 2 displays the AD and its η-relaxation. We can regard η-AAD as a relaxed version of AD in the sense that when $\angle(\boldsymbol{y},\, C) \in (\tau, (1+\eta)\,\tau]$, no demands are placed on the output of the algorithm. This buffer zone allows us to work with a surrogate cone $\widehat{C}$ with much simpler structure, enabling computationally efficient detection. For example, if we form a polyhedral approximation $\widehat{C} = \mathrm{cone}(\widehat{\boldsymbol{A}})$, the distance to $\widehat{C}$ is just the optimal value of the nonnegative least squares problem

$$d(\boldsymbol{y}, \widehat{C}) = \min_{\boldsymbol{x} \geq \boldsymbol{0}} \|\boldsymbol{y} - \widehat{\boldsymbol{A}}\boldsymbol{x}\|_2^2. \quad (5)$$

To implement the angular detector $\mathfrak{D}_\xi^{\widehat{C}}$ for $\widehat{C}$, we just need to solve (5) and compare the optimal value to a threshold. Moreover, it should come with no surprise that whenever $\widehat{C}$ approximates C sufficiently well, we have detector $\mathfrak{D}_\xi^{\widehat{C}} \in \widehat{\mathbb{D}}_{\tau,\eta}^C$, with ξ chosen appropriately.

---

[2] For convex, Lambertian objects, in a point sampling model of image formation, the best known bound on the number of extreme rays in a $\mathcal{V}$-representation of C is quadratic in the number of image pixels: $n = O(m^2)$ [2]. For nonconvex objects or more realistic sampling models, C may not even be polyhedral.

To make this precise, we need a notion of approximation. We will work with the following discrepancy δ:

$$\delta\left(C, \widehat{C}\right) = \max\left\{ \sup_{\boldsymbol{y} \in C, \|\boldsymbol{y}\|=1} d(\boldsymbol{y}, \widehat{C}), \sup_{\boldsymbol{y} \in \widehat{C}, \|\boldsymbol{y}\|=1} d(\boldsymbol{y}, C) \right\} \quad (6)$$

This is just the Hausdorff distance between $C \cap \mathfrak{B}(0, 1)$ and $\widehat{C} \cap \mathfrak{B}(0, 1)$, here $\mathfrak{B}(0, 1)$ denotes a unit ball around 0. Hence, it satisfies the triangle inequality: $\forall \bar{C}$,

$$\delta(C, \widehat{C}) \leq \delta(C, \bar{C}) + \delta(\bar{C}, \widehat{C}). \quad (7)$$

If $\delta(C, \widehat{C})$ is small, we indeed lose little in working with $\widehat{C}$:

Lemma 2.3. *Given cone* C, $\tau > 0$ *and* $\eta \geq 0$, *with* $(1+\eta)\tau \in (0, \frac{\pi}{2})$ *and another cone* $\widehat{C}$, *we have* $\mathfrak{D}_\xi^{\widehat{C}} \in \widehat{\mathbb{D}}_{\tau,\eta}^C$ *whenever*

- $\delta(C, \widehat{C}) \leq \frac{1}{2}\left(\sin(\tau + \eta\tau) - \sin\tau\right)$;
- $\xi \in [\mathrm{asin}(\sin\tau + \delta), \mathrm{asin}(\sin(\tau + \eta\tau) - \delta)]$.

If $\delta(C, \widehat{C})$ is small, then we can simply apply an angular test with cone $\widehat{C}$, and this will implement an approximate angular detector for C.

Methodology. We present a detailed procedure of producing an approximate cone $\widehat{C}$. In Section 3, we will show how to build an ε-approximation $\bar{C} = \mathrm{cone}(\bar{\boldsymbol{A}})$ to C, where $\bar{\boldsymbol{A}} \in \mathbb{R}^{m \times n}$ is a matrix whose columns are images under point illumination. Then in Section 4, via solving a convex optimization problem, we form cone $\widehat{C}$, a γ-approximation to $\bar{C}$, but with much lower complexity. From (7), our resulting cone $\widehat{C}$ $(\varepsilon + \gamma)$-approximates C: $\delta(C, \widehat{C}) \leq \varepsilon + \gamma$.

3. Illumination Cone Approximation

3.1. Extreme Rays of C

As discussed above, under very general circumstances the set of images $\boldsymbol{y}$ form a convex cone. This follows from the linearity of $\boldsymbol{y}[f]$. For the models we will consider, we will see that the linear function $\boldsymbol{y}[\cdot]$ can be written as

$$\boldsymbol{y}[f] = \int_{\boldsymbol{u} \in \mathbb{S}^2} \bar{\boldsymbol{y}}[\boldsymbol{u}]\, f(\boldsymbol{u})\, d\boldsymbol{u}, \quad (8)$$

where $\bar{\boldsymbol{y}} : \mathbb{S}^2 \to \mathbb{R}^m$ is a continuous function. As above, we are interested in $C_0 = \boldsymbol{y}[\mathcal{F}]$. In this case, the vectors $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ form the extreme rays of the cone C_0:

Lemma 3.1. *Suppose that the imaging map* $\boldsymbol{y}$ *satisfies* (8), *with* $\bar{\boldsymbol{y}}[\cdot] : \mathbb{S}^2 \to \mathbb{R}^m$ *continuous. Then for* $C_0 = \boldsymbol{y}[\mathcal{F}]$,

$$\delta\left(C_0, \mathrm{cone}\left(\{\bar{\boldsymbol{y}}[\boldsymbol{u}] \mid \boldsymbol{u} \in \mathbb{S}^2\}\right)\right) = 0. \quad (9)$$

The $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ can be considered images of $\mathcal{O}$ under *point illumination* from direction $\boldsymbol{u}$. With this interpretation, the previous lemma simply asserts that any image $\boldsymbol{y}[f]$ under distant, Riemann integrable illumination f can be arbitrarily well approximated using a conic combination of images

$\bar{\boldsymbol{y}}[\boldsymbol{u}]$ under point illuminations. The conic hull of these extreme images is (up to a set of measure zero) the cone C_0 of images of $\mathcal{O}$ under arbitrary Riemann integrable illumination. We would like a similar expression that works when the ambient level is larger than zero – we would like to also approximate the extreme rays of C_α. The following lemma says that we can use images of the form $\check{\boldsymbol{y}}[\boldsymbol{u}] = \bar{\boldsymbol{y}}[\boldsymbol{u}] + \alpha \boldsymbol{y}_a$, where $\boldsymbol{y}_a$ is the image of $\mathcal{O}$ under ambient illumination:

Lemma 3.2. *Suppose that* $\boldsymbol{y}[f]$ *satisfies* (8) *with* $\bar{\boldsymbol{y}}[\cdot]$ *continuous. Set* $\check{\boldsymbol{y}}[\boldsymbol{u}] = \bar{\boldsymbol{y}}[\boldsymbol{u}] + \alpha \boldsymbol{y}_a$, *with*

$$\boldsymbol{y}_a = \frac{1}{\text{area}(\mathbb{S}^2)} \int_{\boldsymbol{u}} \bar{\boldsymbol{y}}[\boldsymbol{u}] d\boldsymbol{u}, \qquad (10)$$

and $\check{C} = \text{cone}\left(\{\check{\boldsymbol{y}}[\boldsymbol{u}] \mid \boldsymbol{u} \in \mathbb{S}^2\} \cup \{\boldsymbol{y}_a\}\right)$. *Then, we have* $\delta(C_\alpha, \check{C}) = 0$.

So, to work with C_α, we can simply work with a modified set of extreme images $\check{\boldsymbol{y}}[\boldsymbol{u}]$, which are sums of images under point illumination and the ambient image $\boldsymbol{y}_a$. We still need to build a computationally tractable representation for C_α. A natural approach to is to discretize the set of illumination directions, by choosing a finite set $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$. The following lemma asserts that as long as the $\bar{\boldsymbol{y}}[\boldsymbol{u}_i]$ can approximate any point illumination $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ in an absolute sense, the cone generated by the finite set and the cone C_α will not differ too much:

Lemma 3.3. *Let* $\bar{C} = \text{cone}\left(\check{\boldsymbol{y}}[\boldsymbol{u}_1], \ldots, \check{\boldsymbol{y}}[\boldsymbol{u}_N], \boldsymbol{y}_a\right)$, *and*

$$\delta(\check{C}, \bar{C}) \le \frac{2 \sup_{\boldsymbol{u} \in \mathbb{S}^2} \min_i \|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}_i]\|_2}{\eta_\star \alpha \|\boldsymbol{y}_a\|_2}. \qquad (11)$$

Here $\eta_\star = \sup_{\|\boldsymbol{w}\|_2 \le 1} \inf_{\boldsymbol{u}} \left\langle \boldsymbol{w}, \frac{\check{\boldsymbol{y}}[\boldsymbol{u}]}{\|\check{\boldsymbol{y}}[\boldsymbol{u}]\|_2} \right\rangle \ge m^{-1/2}$ *measures the angular spread of* C_α.

This substantially simplifies the problem of approximating C_α: to control the error over *all* possible images, it is enough to control the error over images under *point illumination*. Below, we will see that this is possible, even for nonconvex objects, provided the object is Lambertian.
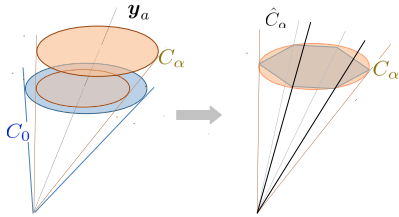

Figure 3. **Cone Approximation:** C_α and its $\mathcal{V}$-approximation.

3.2. Convexity Defect

We will define two complementary quantities which measure the *nonconvexity* of $\mathcal{O}$:

Ambient visibility is a function of the position $\boldsymbol{x}$ on the object boundary. It is defined as the fraction of directions that are visible from $\boldsymbol{x}$, weighted by $\langle \boldsymbol{n}(\boldsymbol{x}), \boldsymbol{u} \rangle$:

$$\tilde{\nu}(\boldsymbol{x}) \doteq \frac{1}{\pi} \int_{\langle \boldsymbol{u}, \boldsymbol{n}(\boldsymbol{x}) \rangle \ge 0} \langle \boldsymbol{n}(\boldsymbol{x}), \boldsymbol{u} \rangle \, \nu(\boldsymbol{x}, \boldsymbol{u}) \, d\sigma(\boldsymbol{u}) \in [0, 1]. \quad (12)$$

Here, ν is a *point-direction visibility* indicator function: $\nu(\boldsymbol{x}, \boldsymbol{u}) = 1$ if point $\boldsymbol{x} \in \partial\mathcal{O}$ can be viewed from direction $\boldsymbol{u} \in \mathbb{S}^2$ and $\nu(\boldsymbol{x}, \boldsymbol{u}) = 0$ otherwise.
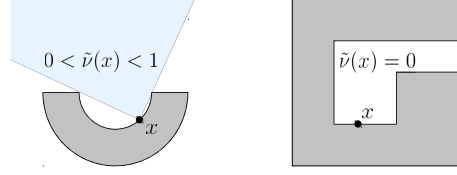

Figure 4. **Ambient Visibility** $\tilde{\nu}(\boldsymbol{x})$

Gnomon length is a function of illumination direction $\boldsymbol{u}$. It is defined as the total length of the edges that cast shadows on the object itself under illumination from $\boldsymbol{u}$.[3]
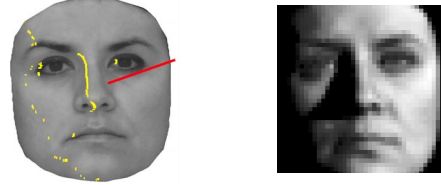

Figure 5. **Shadowing Edges** $\chi[\boldsymbol{u}]$ (yellow) under directional illumination $\boldsymbol{u}$ (red), with corresponding image on the right.

For convex object, $\tilde{\nu}(\boldsymbol{x}) = 1$ and $\chi[\boldsymbol{u}] = 0$ always hold for any point $\boldsymbol{x} \in \partial\mathcal{O}$ and any illumination direction $\boldsymbol{u} \in \mathbb{S}^2$. For more general objects, their nonconvexity will be phrased in terms of the extreme values of these quantities:

Minimum visibility: $\qquad \nu_\star \doteq \inf_{\boldsymbol{x} \in \partial\mathcal{O}} \tilde{\nu}(\boldsymbol{x}) \ge 0$;

Max. gnomon length: $\qquad \chi_\star \doteq \sup_{\boldsymbol{u} \in \mathbb{S}^2} \text{length}\left(\chi[\boldsymbol{u}]\right)$.

Clearly, $\nu_\star = 1$ and $\chi_\star = 0$ for convex objects. For general objects, $1 - \nu_\star$ and $\chi_\star$ can be interpreted as measures of nonconvexity. These two quantities capture complementary information: $\nu_\star$ is localized, depending on properties of the object at a point, while $\chi_\star$ depends more strongly on the global geometry. They will play an important role in building an approximation to the original illumination cone.

3.3. Physical Assumptions

We will introduce hypotheses on the object and the image formation process. Under these hypotheses, we obtain rigorous bounds for the error $\|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}']\|_2$ incurred by approximating $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ with $\bar{\boldsymbol{y}}[\boldsymbol{u}']$ illuminated from another direction $\boldsymbol{u}'$.

Object Geometry. Our bounds pertain to *triangulated* objects, whose boundary is a union of finitely many oriented triangles:

---

[3]For a precise definition of $\chi[\boldsymbol{u}]$, please refer to the supplement.

Definition 3.4 (Triangulated object)**.** *We say that $\mathcal{O} \subset \mathbb{R}^3$ is a triangulated object if for some integer N,*

$$\partial\mathcal{O} = \cup_{i=1}^{N}\triangle_i,$$
$$\triangle_i = \mathrm{conv}\left\{\boldsymbol{v}_i^{(1)}, \boldsymbol{v}_i^{(2)}, \boldsymbol{v}_i^{(3)}\right\}, \quad \boldsymbol{v}_i^{(1)} \neq \boldsymbol{v}_i^{(2)} \neq \boldsymbol{v}_i^{(3)}$$
$$\triangle_i \cap \triangle_j \subseteq \{\emptyset\} \cup \mathcal{V} \cup \mathcal{E}, \qquad \forall\, i \neq j,$$

where $\mathcal{V}$ and $\mathcal{E}$ are sets of vertices and edges respectively, and each face $\triangle_i$ has a unique outward normal $\boldsymbol{n}_i \in \mathbb{S}^2$.

This geometric assumption captures most of the object models that are interesting for computer graphics. Notice that N above can be arbitrarily large – and hence this model can approximate smooth objects.

Object Reflectance. We consider a Lambertian reflectance model. In this model, the object is fully described by its geometry and its *albedo* $\rho : \partial\mathcal{O} \to (0, 1]$, which is the fraction of incoming light that is reflected at each point $\boldsymbol{x} \in \partial\mathcal{O}$. In the Lambertian model, the key quantity linking the illumination f and the image $\boldsymbol{y}$ is the outgoing radiance (radiosity) $g : \partial\mathcal{O} \to \mathbb{R}_+$ at each point $\boldsymbol{x} \in \partial\mathcal{O}$. The map from distant illumination f to outgoing radiance g can be described in terms of two operators:

The ***direct illumination*** operator $\mathcal{D}: \mathcal{F} \to L^2[\partial\mathcal{O}]$ describes the object's reflectance after the first bounce of light from illumination function $f(\boldsymbol{u})$:

$$\mathcal{D}\left[f\right](\boldsymbol{x}) = \int_{\boldsymbol{u}} \bar{\mathcal{D}}[\boldsymbol{u}](\boldsymbol{x}) f(\boldsymbol{u}) d\boldsymbol{u}. \tag{13}$$

For Lambertian objects,

$$\bar{\mathcal{D}}[\boldsymbol{u}](\boldsymbol{x}) = \rho(\boldsymbol{x})\langle \boldsymbol{n}(\boldsymbol{x}), \boldsymbol{u}\rangle_+ \nu(\boldsymbol{x}, \boldsymbol{u}). \tag{14}$$

Here, ν is the same point-direction visibility indicator function in equation (12).

The ***interreflection*** operator $\mathcal{T}: L^2[\partial\mathcal{O}] \to L^2[\partial\mathcal{O}]$ describes the how light reflected off the object illuminates the object itself again:

$$\mathcal{T}[g](\boldsymbol{x}) = \int_{\boldsymbol{x}'\in\partial\mathcal{O}} \frac{\rho(\boldsymbol{x})}{\pi} g(\boldsymbol{x}') \frac{\langle \boldsymbol{n}_{\boldsymbol{x}'}, \boldsymbol{x}-\boldsymbol{x}'\rangle\langle \boldsymbol{n}_{\boldsymbol{x}}, \boldsymbol{x}'-\boldsymbol{x}\rangle}{\|\boldsymbol{x}-\boldsymbol{x}'\|^4} V_{\boldsymbol{x},\boldsymbol{x}'} d\boldsymbol{x}'.$$

Here, V is a *point-point visibility* indicator function: $V_{\boldsymbol{x},\boldsymbol{x}'} = 1$ if point pairs $(\boldsymbol{x}, \boldsymbol{x}') \in \partial\mathcal{O} \times \partial\mathcal{O}$ are mutually visible and $V_{\boldsymbol{x},\boldsymbol{x}'} = 0$ otherwise.

For all of the models that we consider, we will show the operator norm of $\mathcal{T}$ will be strictly smaller than one, and so the operator $\mathcal{I} - \mathcal{T}$ will be invertible. Under this assumption the outgoing radiance on the surface of the object can be written as a convergent series

$$g[f] = \mathcal{D}\left[f\right] + \mathcal{T}\mathcal{D}\left[f\right] + \mathcal{T}^2\mathcal{D}\left[f\right]\ldots = (\mathcal{I}-\mathcal{T})^{-1}\mathcal{D}\left[f\right].$$

Sensor Model. We consider a perspective camera, with a thin lens model that is commonly adopted in computer vision [11]. Suppose the camera has camera gain γ_c, focal length f_c and lens diameter d_c, and its imaging sensor is composed of m non-overlapping squares $I_i \subset \mathbb{R}^2$ of width s_c, then the value of the i-th pixel is generated by integrating the radiance over region I_i, hence

$$y_i = \mathcal{P}_i[g] \doteq \frac{\gamma_c}{4}\left(\frac{d_c}{f_c}\right)^2 \int_{\boldsymbol{z}\in I_i} g(\mathfrak{p}^{-1}(\boldsymbol{z})) \left\langle \frac{\boldsymbol{z}}{\|\boldsymbol{z}\|_2}, \boldsymbol{e}_3\right\rangle^4 d\mu(\boldsymbol{z}).$$

Here, $\mathfrak{p}$ represents perspective projection; its inverse maps an image point to the corresponding point on $\partial\mathcal{O}$. Combining the expressions for pixels $1\ldots m$, we can describe the image vector $\boldsymbol{y}$ via

$$\boldsymbol{y} = \mathcal{P}[g] = [\mathcal{P}_1[g], \cdots, \mathcal{P}_m[g]]^T \in \mathbb{R}^m. \tag{15}$$

Imaging Operator. As long as $\|\mathcal{T}\| < 1$, the image of $\mathcal{O}$ under point illumination from $\boldsymbol{u}$ is given as follows:

Lemma 3.5. *Under our imaging model, with $\mathcal{P}$ as in (15), $\mathcal{T}$ as in (15) and $\mathcal{D}$ as in (14), we have $\boldsymbol{y}[f] = \int_{\boldsymbol{u}} \bar{\boldsymbol{y}}[\boldsymbol{u}]f(\boldsymbol{u})d\boldsymbol{u}$, with $\bar{\boldsymbol{y}}[\boldsymbol{u}] = \mathcal{P}(\mathcal{I}-\mathcal{T})^{-1}\bar{\mathcal{D}}[\boldsymbol{u}]$.*

Under our hypotheses, $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ is continuous in $\boldsymbol{u}$. From Lemma 3.3, if we can approximate these $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ well, we will have a good approximation for the whole cone.

3.4. Perturbation Bounds

We will show how our assumptions can be used to control $\|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}']\|_2$ in terms of $\|\boldsymbol{u} - \boldsymbol{u}'\|_2$. The relationship between $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ and $\boldsymbol{u}$ obviously depends on detailed properties of $\mathcal{O}$, which have been quantified as two non-convexity measurements $\chi_\star$ and $\nu_\star$. The quantity $\chi_\star$ plays a key role in controlling the direct illumination operator:

Lemma 3.6 (Perturbation of direct illumination)**.** *Suppose that $\mathcal{O}$ is a triangulated object with albedo $\rho(\boldsymbol{x}) \in (0, 1]$. Let $\bar{\mathcal{D}}[\boldsymbol{u}] \in L^2[\partial\mathcal{O}]$ be as in (14). Then for all $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{S}^2$,*

$$\left\|\bar{\mathcal{D}}[\boldsymbol{u}] - \bar{\mathcal{D}}[\boldsymbol{u}']\right\|_{L^2}^2 \leq 2\,\mathrm{area}(\partial\mathcal{O}) \left\|\boldsymbol{u} - \boldsymbol{u}'\right\|_2^2$$
$$+ 32\sqrt{2}\,\mathrm{diam}\,(\mathcal{O})\,\chi_\star \left\|\boldsymbol{u} - \boldsymbol{u}'\right\|_2.$$

Here, $\mathrm{area}(\partial\mathcal{O})$ is the area of the object surface and $\mathrm{diam}\,(\mathcal{O})$ is the diameter of the object.

The first term accounts for continuous changes induced by $\langle \boldsymbol{n}(\boldsymbol{x}), \boldsymbol{u}\rangle_+$, and the second term accounts for nonsmooth changes induced by cast shadows. Though cast shadow could introduce a sharp change of radiance for some points, the area of those points is always small when $\|\boldsymbol{u} - \boldsymbol{u}'\|_2$ is small.[4]

While the other quantity $\nu_\star$ is important in controlling the interreflection operator:

Lemma 3.7. *The operator $\mathcal{T}$ satisfies $\|\mathcal{T}\| \leq 1 - \nu_\star$. If $\nu_\star > 0$, $\|\mathcal{T}\| < 1$, and $\left\|(\mathcal{I} - \mathcal{T})^{-1}\right\| \leq \nu_\star^{-1}$.*

---

[4]If $\mathcal{O}$ is convex, the bound simplifies to $\left\|\bar{\mathcal{D}}[\boldsymbol{u}] - \bar{\mathcal{D}}[\boldsymbol{u}']\right\|_{L^2}^2 \leq \mathrm{area}(\partial\mathcal{O}) \left\|\boldsymbol{u} - \boldsymbol{u}'\right\|_2^2$.

For convex objects, $\nu_\star = 1$, and the interreflection term does not participate in image formation process.

We show how to bound $\|\mathcal{P}\|$ in the appendix. By putting these three bounds together, we can obtain a final result:

Theorem 3.8. *Under our hypothesis, for any pair of illumination directions* $\boldsymbol{u}, \boldsymbol{u}' \in \mathbb{S}^2$ *with* $\|\boldsymbol{u} - \boldsymbol{u}'\|_2 \leq \sqrt{2}$,

$$\|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}']\|_2 \leq 2^{1/4} \frac{\beta f_c s_c}{l \nu_\star} \times$$

$$\left(2\mathrm{area}(\partial\mathcal{O}) \|\boldsymbol{u} - \boldsymbol{u}'\|_2^2 + 32\sqrt{2} \,\mathrm{diam}\,(\mathcal{O}) \,\chi_\star \|\boldsymbol{u} - \boldsymbol{u}'\|_2 \right)^{1/2}.$$

Here, $l = \min\{\langle \boldsymbol{e}_3, \boldsymbol{x} \rangle \mid \boldsymbol{x} \in \mathcal{O}\}$ is the depth of the object, and $\beta = \frac{\gamma_c d_c^2}{4 f_c^2}$ is a camera parameter.

Number of Sample Images. To our knowledge, this result, and in particular the perturbation bound Lemma 3.6 are new, and could be useful for other problems in vision and graphics. This bound depends only on properties of the object and imaging system that can be known or estimated. In conjunction with Lemma 3.3, it gives a guideline for choosing the sampling density that *guarantees* a representation that works for every illumination $f \in \mathcal{F}_\alpha$.

In particular, as $\|\boldsymbol{u} - \boldsymbol{u}'\|_2 \to 0$, Theorem 3.8 implies that $\|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}']\|_2$ is proportional to $\|\boldsymbol{u} - \boldsymbol{u}'\|_2^{1/2}$. We can deduce that for ε-approximate guaranteed verification with ambient illumination level α, it would be enough to have

$$n(\alpha, \varepsilon) = \frac{\mathrm{const}(\texttt{sensor}, \texttt{object})}{(\alpha\varepsilon)^4} \qquad (16)$$

sample images – polynomial in the approximation error ε and dimension m. This is possible due to the very special structure of the extreme rays $\bar{\boldsymbol{y}}[\boldsymbol{u}]$ of C_α. In contrast, general cone approximation in $\mathbb{R}^m$ requires a number of samples exponential in m [3].

4. Cone Preserving Complexity Reduction

Although the sample complexity $n(\alpha, \varepsilon)$ in (16) is polynomial in ε^{-1}, it can still be very large. This makes working directly with the dictionary $\bar{\boldsymbol{A}} \in \mathbb{R}^{m \times n}$ problematic in practice. Hence, we want to work with a surrogate $\widehat{\boldsymbol{A}}$ which enables efficient numerical computations while still belongs to the set Ω_0 for guaranteed verification

$$\Omega_0 \doteq \left\{ \widehat{\boldsymbol{A}} \in \mathbb{R}^{m \times n} \mid \delta\left(\mathrm{cone}(\bar{\boldsymbol{A}}), \mathrm{cone}(\widehat{\boldsymbol{A}})\right) \leq \gamma \right\} \quad (17)$$

If $\widehat{\boldsymbol{A}}$ can be expressed as $\boldsymbol{L} + \boldsymbol{S}$, where $\boldsymbol{L}$ has rank r and $\boldsymbol{S}$ has k nonzero entries, product $\widehat{\boldsymbol{A}}\boldsymbol{x}$ can be computed in time $O((m+n)r + k)$, much smaller than $O(mn)$. This model also fits image under varying illuminations (Figure 6): the low-rank term captures the smooth variations [1], while cast shadows are often sparse [21]. The effectiveness of such model under has been noted, e.g., in [4], and exploited for robust photometric stereo by [22].
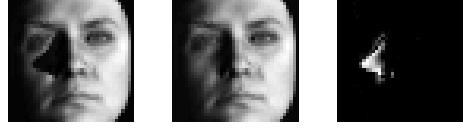

Figure 6. **Low Rank + Sparse Decomposition.** Left: input $\bar{\boldsymbol{A}}$. Middle: low-rank term $\boldsymbol{L}$. Right: sparse term $\boldsymbol{S}$.

To build a framework for complexity reduction with guaranteed approximation quality, we start with:

$$\min_{\boldsymbol{L}, \boldsymbol{S}} \quad \mathrm{rank}\,(\boldsymbol{L}) + \lambda\|\boldsymbol{S}\|_0 \quad \mathrm{s.t.} \quad \boldsymbol{L} + \boldsymbol{S} = \widehat{\boldsymbol{A}} \in \Omega_0.$$

The nonconvex objective rank and ℓ^0-norm can be replaced by nuclear norm and the ℓ^1-norm. The nonconvex domain Ω_0 can be replaced by a convex subset Ω_1:

Lemma 4.1. *If* $\gamma' \leq \frac{\gamma}{\gamma+1}$, *we have* $\Omega_1 \subseteq \Omega_0$, *where*

$$\Omega_1 \doteq \left\{ \widehat{\boldsymbol{A}} \mid \max_{\boldsymbol{x} \geq \boldsymbol{0}, \|\bar{\boldsymbol{A}}\boldsymbol{x}\|_2 \leq 1} \|\bar{\boldsymbol{A}}\boldsymbol{x} - \widehat{\boldsymbol{A}}\boldsymbol{x}\|_2 \leq \gamma' \right\}. \quad (18)$$

Finally, by lifting $\left(\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T\right)$ and duality, we can obtain:

Theorem 4.2. *Let* $(\boldsymbol{L}^\star, \boldsymbol{S}^\star)$ *solve,*

$$\min_{(\boldsymbol{L}, \boldsymbol{S}, \boldsymbol{\mu})} \quad \|\boldsymbol{L}\|_* + \lambda\|\boldsymbol{S}\|_1$$

$$\mathrm{s.t.} \quad \begin{bmatrix} \boldsymbol{I} & \boldsymbol{L}+\boldsymbol{S}-\bar{\boldsymbol{A}} \\ (\boldsymbol{L}+\boldsymbol{S}-\bar{\boldsymbol{A}})^T & \bar{\gamma}\bar{\boldsymbol{A}}^T\bar{\boldsymbol{A}}-\boldsymbol{\mu} \end{bmatrix} \succeq \boldsymbol{0}, \; \boldsymbol{\mu} \geq \boldsymbol{0}. \;(19)$$

with $\bar{\gamma} \leq \left(\frac{\gamma}{1+\gamma}\right)^2$, *then* $\delta(\mathrm{cone}(\bar{\boldsymbol{A}}), \mathrm{cone}(\boldsymbol{L}^\star + \boldsymbol{S}^\star)) \leq \gamma$.

In contrast to existing matrix decompositions (e.g., [4]), which aim at statistical estimation, and measure quality of approximation in Frobenius norm, we guarantee approximation in Hausdorff distance $\delta(\cdot, \cdot)$. This is precisely the measure required for worst case verification. We call (19) a *cone-preserving low-rank and sparse decomposition*. It can be computed efficiently using the Linearized Alternating Direction Method of Multipliers (L-ADMM)([23]), which converges globally with rate $O\,(1/k)$ [10].

5. Numerical Experiment

We render images from 3D triangulated object models following a simplified imaging process $\boldsymbol{y}[f] = \mathcal{P}\mathcal{D}[f]$. Thus, our simulations include cast shadows, but not interreflection. Camera parameters $\gamma_c = f_c = d_c = 1$ and $s_c = 0.003$ are fixed through out our experiments.

Verifying the Perturbation Bound. We consider point illuminations spaced every $\frac{\pi}{360}$ angle in spherical coordinates. The ratio[5]

$$ratio = \max_{\boldsymbol{u}, \boldsymbol{u}'} \frac{\|\bar{\boldsymbol{y}}[\boldsymbol{u}] - \bar{\boldsymbol{y}}[\boldsymbol{u}']\|_2}{\mathrm{PerturbationBound}(\boldsymbol{u}, \boldsymbol{u}')} \qquad (20)$$

is always bounded by 1 for any pair of adjacent $\boldsymbol{u}$ and $\boldsymbol{u}'$. Table 1 verifies this for three different object shapes:

---

[5]Here, PerturbationBound$(\boldsymbol{u}, \boldsymbol{u}')$ is the bound in Theorem 3.8.

Object	Vase	Face	Bunny
ratio	0.1809	0.0302	0.0290

Table 1. **Tightness of the bounds.** Largest ratio r between experimental observation and theoretical upper bound for three different objects. The bound holds in all cases, and is tightest for the vase.

Order of Perturbation Bound. We next consider the behavior of our bounds when $\|\boldsymbol{u} - \boldsymbol{u}'\|_2 \to 0$. Our bounds predict that in the worst case, the change in radiance $\bar{\mathcal{D}}[\boldsymbol{u}]$ should be proportional to $\|\boldsymbol{u} - \boldsymbol{u}'\|_2^{1/2}$. We investigate this using a toy object composed of two perpendicular surfaces S_1 and S_2 shown in Figure 7 with $\boldsymbol{u}$ fixed, and $\boldsymbol{u}'$ changing slowly. Figure 7 (right) shows how $\left\| \bar{\mathcal{D}}[\boldsymbol{u}'] - \bar{\mathcal{D}}[\boldsymbol{u}] \right\|$ depends on $\|\boldsymbol{u} - \boldsymbol{u}'\|_2$. We can see that indeed, both the theoretical prediction and the computed value are proportional to $\|\boldsymbol{u} - \boldsymbol{u}'\|_2^{1/2}$.[6] This suggests that in the worst case, our theory may be tight up to constant factors.
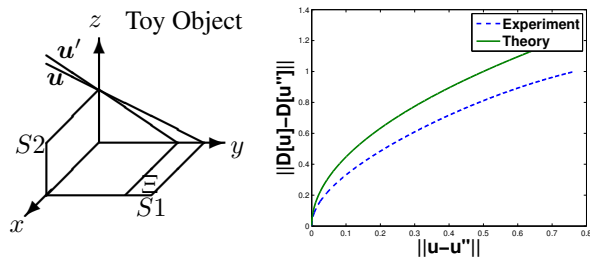

Figure 7. **Order of Perturbation Bound**. Here, in both theory and simulation the change in $\bar{\mathcal{D}}[\boldsymbol{u}]$ is proportional to $\|\boldsymbol{u} - \boldsymbol{u}'\|^{1/2}$.

Cone Preserving Complexity Reduction. We demonstrate the ability of our solution to (19) to reduce the complexity of the representation, while preserving the conic hull. We start $n = 648$ images of a face under point illumination, with resolution 40×40. We solve the low-rank and sparse cone approximation problem in Theorem 4.2 for varying cone distances γ. Figure 8 plots the ratio complexity of $\widehat{\boldsymbol{A}}$ and $\boldsymbol{A}$, or $\frac{(m+n)r+s}{mn}$, where r is the rank of the recovered low-rank term and s is the number of nonzero entries in the recovered sparse term. The decomposition reduces the complexity in all cases; the reduction becomes more pronounced as α increases. This is expected, since the cone C_α becomes smaller as α increases.

Application Sketch. To conceptually verify the advantage of our cone approximation methodology in verification under poor illumination conditions, we compare the receiver operating characteristic (ROC) curves for 5 detection dictionaries obtained from the same 3D face model under ambient level $\alpha = 0.1$: convex cone C_1 composed of 2592

---

⁶In Figure 7, we rescale the theoretical prediction for clearer comparison – our goal is only to show the that the exponent is $1/2$.
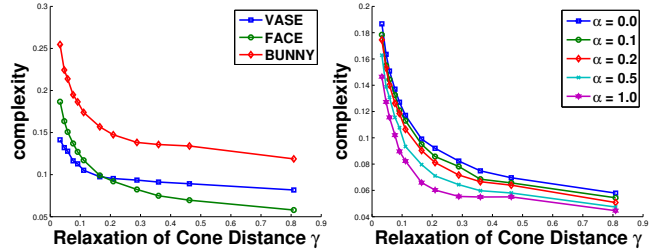

Figure 8. **Cone Complexity Reduction** for different nonconvex objects under zero ambient level ($\alpha = 0$) (left) and for *face* under different ambient illumination levels (right).

images, corresponds to the ε-approximate of the original illumination cone; C_2 is the γ-approximation ($\gamma = 0.11$) of C_1 with $L + S$ structure, whose rank $r = 25$, number of nonzero entries $s = 52699$ and the complexity reduction ratio is 0.0539; C_3 is rendered under 19 illumination directions corresponding to subsets 1 and 2 of Yale B [9] (roughly, the setting of [21]); C_4 is rendered under all 64 illumination directions considered in [9]. Finally, motivated by [1], we also consider the subspace S spanned by 9 principal components of C_1. Also, we have two more ROC curves for verification results based on local gradient features, LBP and HOG respectively: the distances are calculated against the descriptor for ambient image, which act as the training dictionary here.
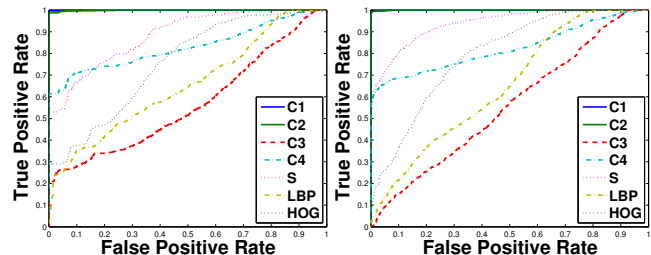

Figure 9. **ROC Curves** for different dictionaries, with test images under uniform random illumination (left) and extreme illumination (right). The dictionaries are C_1: ε approximation, C_2: low-rank and sparse approximation, C_3-C_4: point illuminations distributed similar to [9], S: nine-dimensional linear subspace.

Our test data consist of 1000 positive images under 1000 illumination directions and 3000 negative images of 3 other subjects. We consider two distributions for the illumination directions – uniform on the sphere (roughly corresponding to the "average case"), and uniform on the set of $\boldsymbol{u} \in \mathbb{S}^2$ for which $-0.1 \le u_3 \le 0.4$. Here, the u_3 axis is the camera axis. Arguably, the second set is more challenging. Figure 9 shows the ROC for a simple detection test based on the distance to the models. As suggested by our theoretical results, both C_1 and C_2 perform almost perfectly. The simpler models C_3, C_4, S perform better than chance, but still break down frequently. We view this result as illustrating a trade-off in illumination representation: uniformly good performance is possible, *if* we can afford a more complex repre-

sentation. The cone preserving low-rank and sparse decomposition gives a way to control the complexity of the representation, while still maintaining this good performance.

6. Discussion

There are several directions for future work. Although our cone construction scheme guarantees worst case detection, the number of samples required is likely to be very large, in particular for small ε: when $\varepsilon = 0.01$, our theorem requires about 10^{25} images under ambient level $\alpha = 1$. We believe there should exist a simpler representation if we could take advantage of the structure of shadows.

To use the results in a practical recognition system, we need to account for variations in object pose as well. This can be done using local optimization heuristics, or simply building models at a set of reference poses [9]. It will be important to have very concise models for each pose; the complexity reduction by convex programming is one means of achieving this.

Here, we have considered object instance *verification*, rather than object instance *recognition*. The "yes/no"' question in verification forces us to confront basic questions about the set of images of the object. Nevertheless, we believe our methodology will be useful for recognition as well. For example, one could build models $\widehat{C}$ for each class and assigning the test sample to the closest model in angle. For recognition problems, the formulations and goals for sampling and complexity reduction may also change.

We anticipate three classes of practical application of our results. The first is in instance detection/recognition using 3D models and 2D test images. The second is in instance detection/recognition with active acquisition of training data, e.g., in face recognition for access control [19]. The final, more speculative application is in instance detection/recognition with large families of objects with similar gross shape and appearance. In face recognition, learned models for physical variabilities (albedo and illumination) are often used in conjunction with deformable models [7]. In many practical settings, this approach mitigates the difficulties associated with small training datasets – they can work with as few as one image [20]. Our results could give a way of learning a set of canonical illumination models, which capture effects such as cast shadows, and which could be adapted to each new input subject.

References

[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *PAMI*, 25:218–233, 2003. 1, 2, 6, 7

[2] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions? *IJCV*, 28:270–277, 1998. 1, 2, 3

[3] E. Bronshteyn and D. Ivanov. The approximation of convex sets by polyhedra. *Siberian Math. J.*, 1976. 6

[4] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *JACM*, 58(7), May 2011. 2, 6

[5] V. Chandrasekaran, S. Sanghavi, P. Parillo, and A. Wilsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011. 2

[6] T. Chen, W. Yin, X. S. Zhou, D. Domaniciu, , and T. Huang. Illumination normalization for face recognition and uneven background correction using total variation based image models. In *CVPR*, 2005. 1

[7] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *PAMI*, 23:681–685, 2001. 8

[8] D. Frolova, D. Simakov, and R. Basri. Accuracy of spherical harmonic approximations for images of lambertian objects under far and near lighting. In *ECCV*, 2004. 1, 2

[9] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *PAMI*, 23(6):643–660, June 2001. 1, 7, 8

[10] B. He and X. Yuan. On the o(1/n) convergence rate of the douglas-rachford alternating direction method. *SIAM Journal on Numerical Analysis*, 50(2):700–709, 2012. 6

[11] B. K. Horn. *Robot Vision*. 1st edition, 1986. 5

[12] K.-C. Lee, J. Ho, and D. J. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *PAMI*, 27(5):684–698, May 2005. 1

[13] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, Nov. 2004. 1

[14] X. Mei, H. Ling, and D. Jacobs. Sparse representation of cast shadows via l1-regularized least squares. In *ICCV*, 2009. 1

[15] R. Ramamoorthi. Analytic pca construction for theoretical analysis of lighting variability in images of a lambertian object. *PAMI*, 24(10):1322–1333, Oct. 2002. 1, 2

[16] R. Ramamoorthi, M. Koudelka, and P. Belhumeur. A fourier theory for cast shadows. *PAMI*, 27(2), 2005. 1

[17] A. Savran, B. Sankur, and M. Taha Bilge. Comparative evaluation of 3d vs. 2d modality for automatic detection of facial action units. *Pattern Recognition*, 45(2):767–782, 2012. 2

[18] A. Shashua and T. Riklin-Raviv. The quotient image: Class-based re-rendering and recognition with varying illuminations. *PAMI*, 23:129–139, 2001. 1

[19] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma. Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *PAMI*, 34(2):372–386, 2012. 1, 8

[20] Y. Wang, L. Zhang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras. Face relighting from a single image under arbitrary unknown lighting conditions. *PAMI*, 31(11), 2009. 1, 8

[21] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *PAMI*, 31(2), 2009. 1, 2, 6, 7

[22] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *ACCV*, 2010. 2, 6

[23] X. Zhang, M. Burger, X. Bresson, and S. Osher. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SJIS*, 3(3):253–276, 2010. 6